John von Neumann Institute for Computing

A Topological Model of a "Jumping Gene" Machine

J. Navarra-Madsen

published in

From Computational Biophysics to Systems Biology (CBSB08),
Proceedings of the NIC Workshop 2008,
Ulrich H. E. Hansmann, Jan H. Meinke, Sandipan Mohanty,
Walter Nadler, Olav Zimmermann (Editors),
John von Neumann Institute for Computing, Jülich,
NIC Series, Vol. **40**, ISBN 978-3-9810843-6-8, pp. 125-128, 2008.

© 2008 by John von Neumann Institute for Computing
Permission to make digital or hard copies of portions of this work for personal or classroom use is granted provided that the copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise requires prior specific permission by the publisher mentioned above.

http://www.fz-juelich.de/nic-series/volume40

A Topological Model of a "Jumping Gene" Machine

Junalyn Navarra-Madsen

Department of Mathematics and Computer Science,
Texas Woman's University, Denton TX 76210, USA
E-mail: jnavarramadsen@mail.twu.edu

Mu is similar to a human immunodeficiency virus (HIV) capable of transposing or integrating itself into the host genome. This paper focuses on a topological model describing the changes in 2-dimensional conformation of the DNA upon Mu binding and catalysis.

1 Introduction

Most of the time, due to the difficulty of making *in vivo* experiments, molecular biologists start with an *in vitro* system of a certain biological process they want to understand. One such important process is DNA transposition. Mu virus was the first transposition system studied in vitro[2]. This in vitro transposition system aided in understanding the structure and function of transposition proteins in nucleoprotein complexes called transpososomes[2,1]. The protein-DNA core of the Mu transpososome is composed of a tetramer of the transposase, Mu A, bound to the two special DNA ends called transposons, sometimes referred to as mobile elements. The reasons for constructing a topological model of Mu-DNA complex are two-fold: first, having an in vitro system is not good enough since the process of isolation and storage of a cellular system may change part or the whole system itself; second, there is no known high-resolution structure of the Mu transpososome to date. A mathematical model (topological or geometric) can be helpful in removing redundancies thereby reducing the number of biologically plausible cases and saving valuable time.

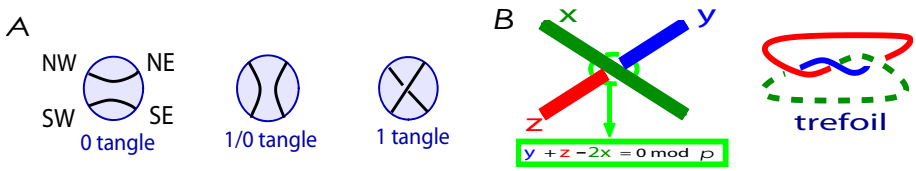

Figure 1. A) Examples of 2-string tangles; B) Definition of colorability. A coloring of diagram of a knot or link or tangle is a function $C : \{arcs\ of\ a\ diagram\} \mapsto \mathbb{Z}_p$ such that at each crossing the relation $y + z - 2x = 0$ mod p holds, where x is the color on the over arc and y and z are the two colors of two under arcs. A three-crossing knot or a trefoil is 3-colorable.

2 Biological Background

In order to topologically study the mechanism of a certain enzyme, molecular biologists start with a circular DNA. Conformational changes on linear DNA made by the enzyme

during catalysis can easily slip and be lost. If one starts with an unknotted double-stranded DNA and obtains a knotted circular DNA, then the enzyme must have acted on the DNA. One possible sequence of enzyme action is: cut, exchange and reseal two DNA segments. Sumners, Ernst and collaborators showed the first mathematical model of DNA recombination using tangles[3]. A n-string tangle is a set of n strings properly embedded in a 3-dimensional ball, Fig. 1A. For a thorough review of tangles and tangle calculus, see a couple of papers by Ernst[5,6]. The 3-dimensional ball represents the protein and the supercoiled string represents double-stranded DNA. Pathania *et al* used circular DNA with properly placed recombination (lox P) and transposition (attL and attR) sites so that they could observe specific changes in these special DNA sites upon protein binding and catalysis[1]. See Fig. 2A. When translated to tangles' parlance, there are two tangles involved, the tangle modeling a better-understood process of Cre recombination[7] and the tangle describing Mu transposition. In the lower right hand corner of Fig. 2B, Cre recombinase changes a 0-tangle to a $\frac{1}{0}$-tangle. This Cre recombination tangle was used as a tool to understand what DNA conformation was trapped in the Mu transpososome upon Mu binding, and a 2-dimensional solution (a tangle with five crossings) obtained by Pathania *et al* after performing all the experiments with different maps where they assumed that the DNA trapped during Mu transposition is *branched supercoiled*, and an example of a branched supercoiled tangle. Most mathematicians are concerned with the question of uniqueness of solution. In the following section, one topological confirmation of uniqueness of this five-crossing solution presented by the in vitro experiments of Pathania *et al* is shown. One strength of this topological modeling is that there is no assumption made about the transposition tangle.

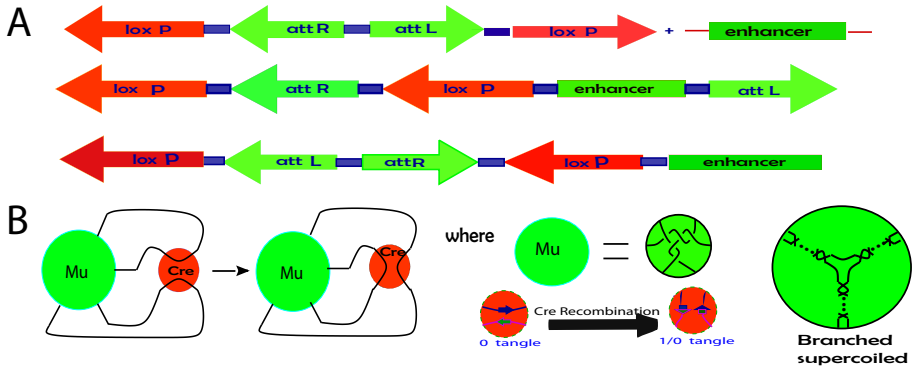

Figure 2. A) Different kinds of maps of Cre recombination and Mu transposition sites; B) 2-dimensional model of the 3-dimensional Mu-DNA transpososome. Pathania *et al* obtained this five-crossing model of Mu-DNA complex[1].

3 Topological Invariant: Colorability

One can use a knot invariant called colorability (Fig. 1B) to mathematically describe the DNA conformational changes formed after the protein binds the DNA and has acted on

it. Darcy *et al* proposed a computational method using colorability to encode DNA conformational changes[8]. Every knot diagram **K** with k crossings has exactly k arcs. By the definition of colorability, there is a $k \times k$ matrix corresponding to a knot diagram **K** with k crossings and k arcs. Therefore, **K** can be colored mod p if and only if the corresponding set of equations has a nontrivial mod p solution[9]. Similarly, tangle diagrams can be colored, Fig. 3A.

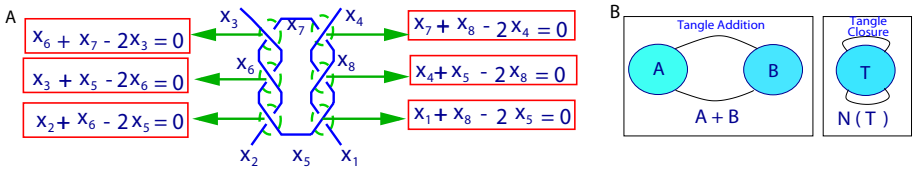

Figure 3. A) An example of coloring a 2-string tangle. There are six crossings (equations) and eight arcs (variables). In[10], it was shown that two tangle invariants can be used to distinguish one n-string tangle from another; B) Tangle addition and tangle closure.

Tangles can be embedded in a knot. Using tangle calculus, a knot can be written as a numerator closure of the sum of two tangles, Fig. 3B. This can then be written as a tangle equation. Two examples of tangle equation is in Fig. 4. One starts with a circular unknot (no crossings when simplified topologically) and ends up with a three crossing knot. These two tangle equations can then be translated into colorability as systems of equations solvable modulo p, $p \in \mathbb{Z}$, given crossings of tangles involved.

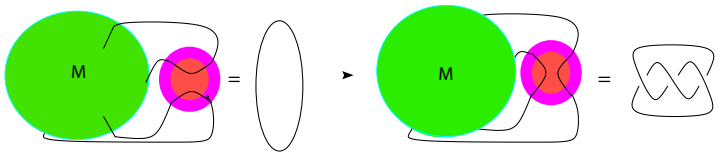

Figure 4. Two examples of tangle equations, one starts with an unknot and ends with a three-crossing knot.

To make the long story short, with the use of two simple tangle invariants found in Ref. 10 and some minor computation, there is one and only one solution to the system of tangle equations, thereby confirming the in vitro solution of Pathania *et al*.

4 Concluding Remarks

One question arises from this study. If given the right length of DNA, will the Mu transpososome take the configurations of crossings greater than five given in the summary table, Table 1.

No. of Crossings	Correct Coloring Matrices	Non-equivalent and Colorable
≤ 4	0	0
5	1	1
6	22	0
7	354	3
8	5019	6

Table 1. Summary table of tangles given the number of crossings in the Mu 2-dimensional configurations.

Acknowledgments

Thank you to the Chancellor's Research Fellows Program of Texas Woman's University for some research support.

References

1. S. Pathania, M. Jayaram, and R. Harshey, *Path of DNA within the Mu transpososome. Transposase interactions bridging two Mu ends and the enhancer trap five DNA supercoils*, Cell **109**, 425–436, 2002.
2. G. Chaconas and R. Harshey, *Transposition of phage Mu DNA, Mobile DNA: Volume II*, (ASM Press, 1752 N Street, NW Washington, DC, 2002).
3. D. W. Sumners, C. Ernst, S. Spengler and N. Cozzarelli, *Analysis of the mechanism of DNA recombination using tangles*, Quarterly Reviews of Biophysics **28**, 253–313, 1995.
4. C. Ernst and D. W. Sumners, *A calculus for rational tangles: applications to DNA recombination*, Math Proc Cambridge Philos Soc **108**, 489–515, 1990.
5. C. Ernst, *Tangle equations*, J Knot Theory Ramifications **5:2**, 145–159, 1996.
6. C. Ernst, *Tangle equations II*, J Knot Theory Ramifications **6**, 1–11, 1997.
7. A. A. Vetcher, A. Y. Lushnikov, J. Navarra-Madsen, R. G. Scharein, Y. L. Lyubchenko, I. K. Darcy, and S. D. Levene, *DNA Topology and Geometry in Flp and Cre Recombination*, J Mol Biol **357**, 1089–1104, 2006.
8. I. K. Darcy, J. Chang, N. Druivenga, C. McKinney, R. Medikonduri, S. Mills, J. Navarra-Madsen, A. Ponnusamy, J. Sweet, and T. Thompson, *Coloring the Mu transpososome*, BMC Bioinformatics **7**, 435–454, 2006, `http://www.biomedcentral.com/1471-2105/7/435`.
9. R. H. Fox, *A Quick Trip Through Knot Theory*, in Topology of 3-manifolds and related topics (edited by M. K. Fort), 120–167, 1961.
10. J. Navarra-Madsen and I. K. Darcy, *Coloring n-string Tangles*, submitted to Journal of Knot Theory and its Ramifications, 2008.